PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-219929
(43)Date of publication of application : 18.08.1995

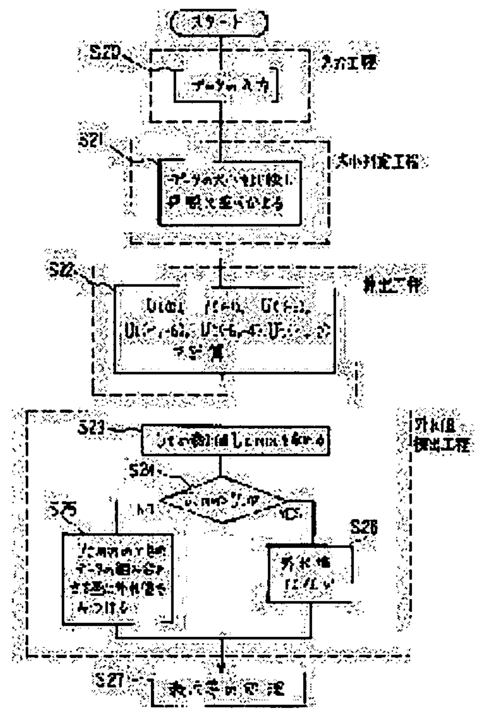
(51)Int.Cl. G06F 17/18
G05B 19/418

(21)Application number : 06-008871
(22)Date of filing : 28.01.1994

(71)Applicant : MITSUBISHI ELECTRIC CORP
(72)Inventor : UEDA TAICHIRO

(54) METHOD FOR DETECTING DEVIATED VALUE AND DATA PROCESSOR

(57)Abstract:
PURPOSE: To find out a deviated value in obtained data (yield amount, reacted amount).
CONSTITUTION: Data are inputted in an input process and the inputted data are ascendingly arranged in a size judging process. Then a detected statistic amount from which the smallest value is excluded is calculated in a calculating process. The detected statistic amount is calculated from the data of combinations to be estimated by similar operation. The detected statistic amount from which no data are removed is also calculated. In a deviated value detecting process, the combination of data minimizing the detected statistic amount is found out and the removed value is regarded as a deviated value. When the detected statistic amount from which 110 data are removed is minimum, a deviated value does not exist.



LEGAL STATUS

[Date of request for examination] 09.10.1997
[Date of sending the examiner's decision of rejection] 03.07.2001
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]
[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

CLAIMS

[Claim(s)]
[Claim 1] The input process which inputs the value of the outlier (method-of-detection a) N individual (N>=3) which has the following processes, (b) The size judging process of judging the size relation of the value of N pieces inputted according to the above-mentioned input process, (c) Based on the size relation judged according to the above-mentioned size judging process, the combination of the value of N pieces and the combination of the value of less than N pieces except the candidate of an outlier are searched for. The blank detection process of detecting an outlier based on the detection statistic computed by the calculation process and the (d) above-mentioned calculation process which compute a detection statistic using a predetermined formula to the combination searched for.
[Claim 2] The above-mentioned calculation process is the outlier method of detection according to claim 1 characterized by computing a detection statistic using two or more combination of the value more than n piece (n=N-s) which continues on the size relation judged according to the size judging process when detecting less than s outliers.
[Claim 3] The outlier method of detection according to claim 1 or 2 characterized by providing the following The above-mentioned outlier detection process is a lowest-selection process which chooses the minimum thing in the detection statistic calculated from the combination of the value of less than N pieces. The outlier judging process which makes an outlier the value which was not included in the combination which computed the selected minimum value when the selected minimum value was smaller than the detection statistic calculated from the combination of the value of N pieces
[Claim 4] It is the outlier method of detection according to claim 1, 2, or 3 which the above-mentioned formula has the 1st item which is in the inclination which becomes small when the candidate of an outlier is removed, and the 2nd item which will become large if the candidate of an outlier is removed, and is characterized by for the above-mentioned calculation process computing the value of the 1st and the 2nd item, and calculating a detection statistic by both sum.
[Claim 5] It is the outlier method of detection according to claim 4 characterized by for the above-mentioned formula having an amendment correction term for the 1st and the 2nd item further in addition to the 1st and 2nd item, and for the above-mentioned calculation process computing the value of the 1st, the 2nd, and the 3rd item, and calculating a detection statistic by three persons' sum.
[Claim 6] The 1st item of the above is the outlier method of detection according to claim 4 or 5 characterized by using distribution of the value of less than N pieces which calculates a detection statistic.
[Claim 7] The 2nd item of the above is the outlier method of detection according to claim 4 or 5 characterized by using the number of the candidate of the outlier in the case of calculating a detection statistic.
[Claim 8] The 2nd item of the above is the outlier method of detection according to claim 7 characterized by using what carried out the multiplication of the predetermined coefficient to the number of the candidate of an outlier.
[Claim 9] It is the outlier method of detection 1 which the above-mentioned formula has the coefficient to the distribution and distribution of less than N pieces of a value which calculate a detection statistic, and is characterized by the above-mentioned calculation process calculating a detection statistic by the multiplication of distribution and a coefficient, 2, or given in three.
[Claim 10] The above-mentioned formula is the claims 1-8 characterized by being created on the basis of the variable selection criterion of regression analysis, or the outlier method of detection given in nine.
[Claim 11] The above-mentioned outlier method of detection is the outlier method of detection according to claim 1 characterized by having the processing process which processes the inputted value between an input process and a size judging process further.
[Claim 12] The above-mentioned processing process is the outlier method of detection according to claim 11 characterized by processing the value which does not depend on time for the value depending on the time inputted according to the input process.

[Claim 13] The above-mentioned processing process is the outlier method of detection according to claim 11 characterized by calculating a leverage from the value inputted according to the input process.
[Claim 14] The above-mentioned processing process is the outlier method of detection according to claim 11 characterized by calculating the data of a regression-analysis model from the value inputted according to the input process.
[Claim 15] The above-mentioned processing process is the outlier method of detection according to claim 11 characterized by calculating the data of a canonical-correlation-analysis model from the value inputted according to the input process.
[Claim 16] The above-mentioned processing process is the outlier method of detection according to claim 11 characterized by calculating the discriminant-function value of each group when the value inputted according to the input process is classified into two or more groups and two or more factors perform discriminant analysis.
[Claim 17] The data processor equipped with an outlier detection means to perform the above-mentioned claims 1-15 or the outlier method of detection given in 16, and to detect an outlier, a measurement means to measure and separate from the value of N pieces and to input into a district prosecutor's office appearance means, and an output means to tell the outlier detected by the outlier detection means.
[Claim 18] The above-mentioned data processor is a data processor according to claim 17 characterized by having a data-processing means to perform predetermined processing further using the remaining value except the outlier detected by the outlier detection means.

[Translation done.]

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Industrial Application] This invention relates to the equipment using the method of detecting the outlier of the data in a production process, quality control, research and development, quality improvement, etc., and its method.
[0002]
[Description of the Prior Art] For example, when measuring with [of a product] a performance rose, or when checking meter of power meter or a water meter etc. and measuring an experimental data further, the data in which the performance of a specification blank is shown, and the data in which unusual measured value is shown exist in the obtained data. Thus, the data of a specification blank and outlying observation are unsuitable data produced from measurement environment or the measuring device itself in many cases. Such unsuitable data are made to call it an outlier below here. Since an outlier is not a value which should originally be measured, the technique of detecting and removing an outlier from various data which were mentioned above has been considered from the former. The method of detecting an outlier from data has a ****** thing in the statistics technique from the former. an outlier is extremely big -- it is -- it is -- it is the thing of the data which take a small value For example, 5.71, 6.57, 7.29, 8.06, 10.00, and 15.00 are considered. If it plots, it will become like drawing 12 . 15.00 seems to be an outlier when drawing 12 is seen. By the statistics technique, an outlier uses the formula which authorizes the outlier at the time of one piece as one piece. The formula for two pieces is used at the time of two pieces. Since it is statistical inspection, it is necessary to decide level of significance (level of significance) beforehand. As level of significance, 5% or 1% is used traditionally. It is shown that the probability which is mistaken when making an outlier and a judgment with 5% of level of significance by statistical inspection is 5%. There is 5% or 1% of table corresponding to the formula, and if the value and table which were calculated by live data are compared and it becomes size, it will consider as an outlier. However, there is 5% as level of significance. Or it will be said that there is 1%.
[0003] Thus, by the conventional statistics technique, when the number of an outlier is one piece and it is two pieces, calculus changes with times of being three pieces, or there are troubles (when it is 5%, even if it concludes it as an outlier, it cannot be said as an outlier at the time of 1%) that a conclusion changes with differences (5%, 1 etc.%, etc.) in level of significance (called level of significance). Moreover, calculus changes with the outlier of a big value, and outliers of a small value. Since it is the statistics technique, 5% and 1% of table is also required.
[0004] Specifically, it explains using drawing 13 and drawing 14 . It is drawing 13 and, for a computer (with FDD), and 3, as for a printer and 5, a display unit and 4 are [1 / an information processor and 2 / a keyboard and 6] floppy disks. The floppy disk 6 by which program routine was memorized is inserted in a computer (with FDD) 2, operation software is driven, and an information processor 1 is started. Program routine will be loaded from a floppy disk 6, and it will be in the input waiting state.
[0005] Drawing 14 is a flow chart for explanation of the conventional example. Step 1 is the stage of inputting data continuously from a keyboard 5. At Step 2, if the number of an outlier is inputted and it becomes one, the statistic which made the small value or the large value the outlier at Step 3 and Step 4 will be calculated, respectively. It is because calculation methods differ in a small value and a large value if it says why it asks separately. Moreover, when the number of an outlier is two pieces, it is Step 7 from Step 5, and when a small value is made into a two-piece outlier, it makes a small value and one large value at a time into an outlier and a large value is made into a two-piece outlier, a statistic is calculated by the respectively different calculation method.
[0006] At Step 8, size comparison of the calculated value which looked at the table and was calculated at the above-mentioned step, and the significant point in a table is carried out. At Step 9, when the calculated value is larger than a significant point, it is recognized as an outlier. Step 10 is not made into an outlier when the calculated value is smaller. A result is indicated at Step 11. Although drawing 14 is the case where level of significance (level of significance) is

5%, if it is 1% of case, 1% of table is needed.
[0007] The example which detects an outlier by the conventional statistics technique is shown. It is referred to as 5.71, 6.57, 7.29, 8.06, 10.00, and 15.00 as data. The formula of the amount of Grubbs official approval and several 1 are used.

[0008]
[Equation 1]
$$T_i = (x_i - \bar{x}) / S$$

ただし $S = \Sigma (x_i - \bar{x}) / (n-1)$, 　　n：サンプル数

[0009] Since 15.00 is considered to be an outlier, it compares with the value which has appeared in the table in quest of the maximum maxTi of Ti (i= 1, 2, ..., n).
It was set to maxTi=1.841. When the table of Grubbs and Table 1 are seen, it is 1.94 at the time of 1.82, a measurement size n= 6, and 1% of level of significance at the time of a measurement size n= 6 and 5% of level of significance. Therefore, it is maxTi>1.82 and maxTi<1.94. Therefore, 15.00 can be called outlier at 5% of level of significance. At 1% of level of significance, it cannot be called an outlier. Thus, a conclusion changes with differences in level of significance.
[0010]
[Table 1]
Ｇｒｕｂｂｓの数表の一部

n	5%	1%
4	1.46	1.49
5	1.67	1.75
6	1.82	1.94
7	1.94	2.10

[0011] (A part is quoted from Vic Barnett and Toby Lewis(1978):"Outliers in Statistical Data" John Wiley& Sons.p.298)
[0012] Next, the example of the masking effect is given. It is referred to as 5.71, 6.57, 7.29, 8.06, 14.80, and 15.00 as data. It is set to maxTi=1.29 by this data. It is maxTi<1.82. Therefore, there will be no outlier. This is an example which does not necessarily detect an outlier, when it is called the masking effect and the candidate of an outlier has 14.80 and two 15.00 as mentioned above, and an outlier is authorized as one by the conventional method.
[0013]
[Problem(s) to be Solved by the Invention] As explained above, in the conventional thing, a calculation method changes with numbers of an outlier. Moreover, there was a trouble that a calculation method changed with character (are they the outlier of the larger one or the outlier of the smaller one?) of an outlier (40 or more sorts of formulas have appeared in Vic Barnett and Toby Lewis(1978):"Outliers in Statistical Data" John Wiley & Sons). Moreover, calculated value and a table need to be size compared. Moreover, there was a trouble that a conclusion changed with differences in level of significance (5%, 1 etc.%, etc.). Moreover, by the conventional method, when it was called the masking effect, for example, there were two candidates of an outlier, when the outlier was authorized as one, there was a trouble of necessarily not detecting an outlier.
[0014] This invention aims at obtaining the outlier method of detection which needs to change a calculation method neither by the number of an outlier, nor the character of an outlier, without being made in order to solve the above troubles, and using a table like before. Moreover, it aims at obtaining the outlier method of detection which can avoid the masking effect. Moreover, when detecting an outlier, computation is easy as much as possible, and computational complexity is also aimed at obtaining the outlier method of detection which there is and ends. [little] Furthermore, it aims at offering the data processor using these outlier methods of detection.
[0015]
[Means for Solving the Problem] The outlier method of detection concerning this invention has the following processes.

(a) The input process which inputs the value of N pieces (N>=3), the size judging process of judging the size relation of the value of N pieces inputted according to the (b) above-mentioned input process, (c) Based on the size relation judged

according to the above-mentioned size judging process, the combination of the value of N pieces and the combination of the value of less than N pieces except the candidate of an outlier are searched for. The blank detection process of detecting an outlier based on measurement, such as detection computed by the calculation process and the (d) above-mentioned calculation process which compute a detection statistic using a predetermined formula to the combination searched for.
[0016] When detecting less than s outliers, the above-mentioned calculation process creates two or more combination of the value more than n piece (n=N-s) which continues on the size relation judged according to the size judging process, and is characterized by computing a detection statistic using such combination.
[0017] The above-mentioned outlier detection process is characterized by having the outlier judging process which makes an outlier the value which was not included in the lowest-selection process which chooses the minimum thing in the detection statistic calculated from the combination of the value of less than N pieces, and the combination which computed the selected minimum value when the selected minimum value was smaller than the detection statistic calculated from the combination of the value of N pieces.
[0018] The above-mentioned formula has the 1st item which is in the inclination which becomes small when the candidate of an outlier is removed, and the 2nd item which will become large if the candidate of an outlier is removed, and the above-mentioned calculation process is characterized by computing the value of the 1st and the 2nd item and calculating a detection statistic by both sum.
[0019] The above-mentioned formula has an amendment correction term for the 1st and the 2nd item further in addition to the 1st and 2nd item, the above-mentioned calculation process computes the value of the 1st, the 2nd, and the 3rd item, and it is characterized by calculating a detection statistic by three persons' sum.
[0020] The 1st item of the above is characterized by using distribution of the value of less than N pieces which calculates a detection statistic.
[0021] The 2nd item of the above is characterized by using the number of the candidate of the outlier in the case of calculating a detection statistic.
[0022] The 2nd item of the above is characterized by using what carried out the multiplication of the predetermined coefficient to the number of the candidate of an outlier.
[0023] The above-mentioned formula has the coefficient to the distribution and distribution of less than N pieces of a value which calculate a detection statistic, and the above-mentioned calculation process is characterized by calculating a detection statistic by the multiplication of distribution and a coefficient.
[0024] The above-mentioned formula is characterized by being created on the basis of the variable selection criterion of regression analysis.
[0025] The above-mentioned outlier method of detection is characterized by having the processing process which processes the inputted value between an input process and a size judging process further.
[0026] The above-mentioned processing process is characterized by processing the value which does not depend on time for the value depending on the time inputted according to the input process.
[0027] The above-mentioned processing process is characterized by calculating a leverage from the value inputted according to the input process.
[0028] The above-mentioned processing process is characterized by calculating the data of a regression-analysis model from the value inputted according to the input process.
[0029] The above-mentioned processing process is characterized by calculating the data of a canonical-correlation-analysis model from the value inputted according to the input process.
[0030] The above-mentioned processing process is characterized for the discriminant-function value of each group by calculation ******, when the value inputted according to the input process is classified into two or more groups and two or more factors perform discriminant analysis.
[0031] Moreover, the data processor concerning this invention is equipped with an outlier detection means to perform the outlier method of detection and to detect an outlier, a measurement means to measure and separate from the value of N pieces and to input into a district prosecutor's office appearance means, and an output means to tell the outlier detected by the outlier detection means.
[0032] The above-mentioned data processor is characterized by having a data-processing means to perform predetermined processing further using the remaining value except the outlier detected by the outlier detection means.
[0033]
[Function] In the 1st invention, if the value of N pieces is inputted according to an input process, the size relation of a value will be judged according to a size judging process, and let some of values of the larger one, or values of the smaller one be the candidates of an outlier. A calculation process computes a detection statistic using a predetermined formula to each combination for which searched for the combination of the value of N pieces, and the combination of

the value of less than N pieces except the candidate of an outlier first, and the degree was asked. An outlier detection process detects an outlier based on the computed detection statistic.

[0034] The calculation process in the 2nd invention calculates a detection statistic by the predetermined formula using the combination of the value followed more than n piece (n=N-s) based on the size of the value judged according to the size judging process. For example, when five pieces (N= 5) are inputted according to an input process and it is going to detect a maximum of two outliers (s= 2), one combination is created using three input values from the larger one. Moreover, another combination is created using four input values from the larger one. Moreover, one combination is created using three middle values except maximum and the minimum value. Moreover, combination is created using three input values of the smaller one, and four input values of the smaller one, respectively.

[0035] The outlier detection process in the 3rd invention chooses first the minimum thing in the detection statistic of the combination of the value of less than N pieces computed by the calculation process. Next, let the value which was not included in the combination which compared the selected minimum value with the detection statistic calculated from the combination of the value of N pieces, and computed the selected minimum value when the minimum value was smaller be an outlier. Moreover, when the minimum value is larger, it judges with an outlier being nothing.

[0036] The formula in the 4th invention has the 1st item which is in the inclination which becomes small when the candidate of an outlier is removed, and the 2nd item which becomes large when the candidate of an outlier is removed, and calculates a detection statistic by both sum. If the value from which it separated most by this formula when there was an outlier is removed, a detection statistic will serve as the minimum.

[0037] In addition to the above 1st and the 2nd item, the formula in the 5th invention has the 3rd item. This 3rd item is an amendment amendment item about the above 1st and the 2nd item. The 1st item, the 2nd item, and the 3rd item are added, and a detection statistic is calculated.

[0038] The formula in the 6th invention includes distribution of the value of less than N pieces which calculates a detection statistic in the 1st item. Therefore, if the value from which it separated most is removed, the value of distribution will become small and the value of the 1st item will become small.

[0039] The formula in the 7th invention contains the number of the candidate of the outlier in the case of calculating a detection statistic in the 2nd item. Therefore, if it is going to detect many number of outliers, the value of the 2nd item will become large.

[0040] What carried out the multiplication of the predetermined coefficient to the 2nd item to the number of the candidate of an outlier is used for the formula in invention of the octavus.

[0041] The formula in the 9th invention carries out the multiplication of the coefficient to distribution of the value of less than N pieces which calculates a detection statistic, and calculates a detection statistic.

[0042] The formula in the 10th invention creates the formula which calculates a detection statistic on the basis of the variable selection criterion of regression analysis.

[0043] In the 11th invention, since the value inputted according to the input process is convertible for the data which can calculate a detection statistic with a processing process, the data of various kinds can be inputted.

[0044] In the 12th invention, the value for which amends an increase or the inclination which decreases in proportion to time from the value inputted according to the input process, and it does not depend on time is processed. And a detection statistic can be computed from the amended value and an outlier can be calculated.

[0045] In the 13th invention, when two or more weighted solidity is in one sample, the diagonal element of a leverage is calculated and an outlier is calculated in quest of a detection statistic based on the calculated value.

[0046] In the 14th invention, when the inputted value can apply the technique of regression analysis, the remainder of regression analysis is calculated and an outlier is calculated in quest of a detection statistic based on the calculated value.

[0047] In the 15th invention, when the inputted value is data of a canonical-correlation-analysis model, canonical correlation analysis is performed and it asks for two synthetic variate functions, and from this, a synthetic variate function value is calculated and an outlier is calculated in quest of a detection statistic based on the value which calculates a leverage from a synthetic variate function value and by which the leverage was calculated.

[0048] In the 16th invention, when the inputted value is classified into two or more groups and two or more factors perform discriminant analysis, in quest of a detection statistic, an outlier is calculated based on the value which calculated the discriminant-function value and was calculated.

[0049] The data processor in the 17th invention measures the value of N pieces by the measurement means, from this measured value, detects an outlier by outlier detection means to perform the above-mentioned outlier method of detection, and tells an outlier by the output means.

[0050] The data processor in the 18th invention measures the value of N pieces by the measurement means, and performs predetermined processing from this measured value using the remaining value except the outlier which

detected the outlier and was detected by the data-processing means by outlier detection means to perform the above-mentioned outlier method of detection.

[0051]

[Example]

Drawing 13 explained in the example 1. conventional example is explained as drawing for explaining the equipment of this example again. It is drawing 13 and, for a computer (with FDD), and 3, as for a printer and 5, a display unit and 4 are [1 / an information processor and 2 / a keyboard and 6] floppy disks. The hardware composition of this invention does not change with the conventional example, but inserts in a computer (with FDD) 2 the floppy disk 6 by which program routine was memorized, drives operation software, and starts an information processor 1. Program routine will be loaded and it will be in the input waiting state. If data are keyed from a keyboard 5, program routine will operate, and a processing result will be displayed on a display 3, and a processing result will be printed on a printer 4.

[0052] In this example, in order to compute a detection statistic, several 2 is used.

[0053]

[Equation 2]

$$Ut = n \log \sigma + 2S$$

ただし、nはサンプル数

Sは外れ値の候補の個数

σは$x_1, x_2, \text{----} x_n$をサンプルデータとすると、

$$\sigma = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n}}$$

[0054] What is necessary is just to find the combination of the sample from which the value of this statistic becomes the minimum. Drawing 1 is a flow chart for explanation of this invention. Step 20 is an input process which inputs the data from a keyboard 5 continuously. For example, x1, x2, x3, x4, and x5 Five data are inputted. In this case, it will be set to N= 5 if the number of the data to input is set to N. Step 21 compares the size of the inputted data, for example, is x1 <x2 <x3 <x4 <x5 to ascending order. It arranges like. This process is a size judging process.

[0055] Thus, by rearranging data into ascending order, it becomes easy to find the candidate of an outlier. If the number of the candidate of an outlier is set to s (s>=1), the candidate of an outlier will say from the property, and s pieces or a large value, and a small value will both unite him from s pieces from the largest value, and the smallest value, and he will be considered to be s pieces.

[0056] Step 22 is a calculation process which calculates the detection statistic Ut by the formula in this example from these data constellations. In the case of the number s= 1 of the candidate of an outlier, it is (x1, x2, x3, x4, and x5) to x1 first. The detection statistic when removing is calculated. Let this be the detection statistic Ut (-1). It is x5 like the following. The time of removing is set to Ut (-5). In the case of the number s= 2 of the candidate of an outlier, it is x1. x2 It is Ut (-1, -2) about the time of removing. It carries out. x4 x5 It is Ut (-4, -5) about the time of removing. It carries out and is x1. x5 It is Ut (-1, -5) about the time of removing. It carries out and a detection statistic is calculated except for a sample sequentially from a small value or a large value. Thus, the detection statistic of the combination considered is calculated, respectively. The number s of the candidate of an outlier shall be beforehand defined by the system here. Or the number s of the candidate of an outlier shall be specified by the operator or the program. Or the number of the candidate of an outlier presupposes that it is possible to set it as the degree of calculation freely.

[0057] In addition, the number of the candidate of the outlier given here is not the number which must be found as an outlier. The number of the candidate of an outlier who says here means the maximum number detected as an outlier. For example, in the case of the number s= 2 of the candidate of an outlier, it is not meant that the case where a maximum of two outliers are found is said, and the two numbers of an outlier are surely found. Therefore, in the case of the number s= 2 of the candidate of an outlier, when an outlier is zero piece and an outlier is one piece, a result, such as a case where they are two outliers, can be considered. When calling it the number s of the candidate of an outlier like the following, a number detectable as an outlier of maximums shall be shown. Thus, what is necessary is not to set the number of outliers as a specific number, and just to specify the maximum of the number of outliers in this example and example mentioned later.

[0058] Step 23 is the stage of finding the minimum value (Utmin) of the detection statistic Ut. Step 24 is the detection statistic Ut when not removing the candidate of an outlier (0). It is the stage which compares Utmin. Step 25 makes an

outlier the value which was not included in the combination of the data when calculating Utmin, when the Utmin is smaller. Step 26 is Ut (0). By the case where it becomes the minimum, an outlier presupposes "There is nothing" at this time. From Step 23 to the step 26 is an outlier detection process. Step 27 gives an indication etc.

[0059] Next, the flow of this flow chart is explained using data. The following five data are inputted at Step 20.
The data inputted at 5.71, 6.57, 7.29, 8.06, and 13.32 step 21 are rearranged into ascending order. Several 2 is used at Step 22 and the value of Ut is calculated. For example, it will be set to Ut(-1) =5.908 if a detection statistic is expressed as Ut (-1) since it is a measurement size n= 4 and the number s= 1 of the candidate of an outlier, when the smallest value 5.71 is removed. Since the detection statistics when removing the biggest value 13.32 are a measurement size n= 4 and the number s= 1 of the candidate of an outlier, if they express Ut (-5), they will be set to Ut(-5) =1.440. Since it is a measurement size n= 3 and the number s= 2 of the candidate of an outlier when both 5.71 and 13.32 are removed, it is Ut (-1, -5). It will be set to Ut(-1, -5) =2.509 if expressed. moreover, if Ut (-1, -2), Ut (-4, -5), Ut (-1, -2, -5), and Ut (-1, -4, -5) are calculated, and the detection statistic Ut of the number s= 3 of the candidate of an outlier (0) is calculated namely,, it will become as it is shown in Table 2

[0060]

[Table 2]

サンプルの組み合わせと検出統計量の値

外れ値の候補 大きい / 小さい	なし	1 3. 3 2	1 3. 3 2 8. 0 6
なし	4. 9 3 0 (S＝0)	(S＝1) 1. 4 4 0 (最小値)	2. 6 8 9 (S＝2)
5. 7 1	5. 7 1 (S＝1)	2. 5 0 9 (S＝2)	3. 9 5 7
5. 7 1 6. 5 7	6. 9 5 7 (S＝2)	4. 0 9 1	

[0061] When the minimum value Utmin of Ut calculated above at Step 23 is calculated, it is Utmin=Ut(-5) =1.440. Detection statistic Ut when not removing the candidate of an outlier at Step 24 (0) = 4.930 is calculated and they are Utmin and Ut (0). It compares. Then, Ut(-5) >Ut(0) Since it is not materialized, it goes to Step 25 and let combination of the data at the time of Ut (-5) be an outlier. That is, 13.32 turns out to be an outlier. Outputs, such as a display of an outlier 13.32, are performed at Step 27.

[0062] Thus, even if it is the case where the number of the candidate of an outlier is 3, the number of the detected outlier is one and can detect an outlier in the range within the number of the candidate of an outlier.

[0063] In the case of the data of 5.71, 6.57, 7.29, 8.06, and 13.32 shown in this example, it is Takeuchi to make 13.32 into an outlier. Same result is brought at ** (1980) "statistics [in a phenomenon and action] mathematical" **** Co.

[0064] sigma+2s of formula Ut=nlog of the detection statistic shown in several 2 of this example was considered from the analogy of AIC (AKAIKE'S information criterion). n is a measurement size and s is the number of the candidate of an outlier, and sigma2. Distribution and sigma are standard deviation. The 1st term of this formula will tend to become small if the sample which is the candidate of an outlier is removed. Because, distributed sigma2 It is because it will become small if an outlier is removed. Moreover, it is because n will become a value small like 5 to 4 and 5 to 3, for example since it is n=(several data N)- (the number s of the candidate of an outlier), if the number of measurement size n removed as a candidate of an outlier increases. The 2nd term will increase, if a measurement size increases. Therefore, when an outlier is removed, it is possible that the sum Ut of the 1st term and the 2nd term becomes the minimum.

[0065] Next, this relation is described using the data 1 of Grubbs. The data 1 of Grubbs are shown below.
2.02, 2.22, 3.04, 3.23, 3.59, 3.73, 3.94, 4.05, 4.11, and the 4.13 total number of data are ten pieces. The value which calculated nlogsigma for the value which calculated logsigma from this data to Table 3 is shown in Table 4, and the value which calculated sigma+2s of Ut=nlog is shown in Table 5.

[0066]

[Table 3]

Ｇｒｕｂｂｓのデータ１によるｌｏｇσの値

外れ値の候補 大＼小	なし	４．１３	４．１３ ４．１１	４．１３ ４．１１ ４．０５
なし	−０．３１３	−０．３１８	−０．３３７	−０．３７６
２．０２	−０．５１４	−０．５１６	−０．５３３	——
２．０２ ２．２２	−０．９５	−０．９６７	——	——
２．０２ ２．２２ ３．０４	−１．１８３	——	——	——

[0067]
[Table 4]

Ｇｒｕｂｂｓのデータ１によるｎｌｏｇσの値

外れ値の候補 大＼小	なし	４．１３	４．１３ ４．１１	４．１３ ４．１１ ４．０５
なし	−３．１２７	−２．８５８	−２．６９７	−２．６２９
２．０２	−４．６３０	−４．１２７	−３．７３２	——
２．０２ ２．２２	−７．６	−６．７７１	——	——
２．０２ ２．２２ ３．０４	−８．２８４	——	——	——

[0068]
[Table 5]

Ｇｒｕｂｂｓのデータ１によるＵｔ＝ｌｏｇσ+２Ｓの値

外れ値の候補 大きい＼小さい	なし	４．１３	４．１３ ４．１１	４．１３ ４．１１ ４．０５
なし	(Ｓ＝０) −３．１２７	(Ｓ＝１) −０．８５８	(Ｓ＝２) １．３０３	(Ｓ＝３) ３．３７１
２．０２	(Ｓ＝１) −２．６３０	(Ｓ＝２) −０．１２７	(Ｓ＝３) ２．２６８	——
２．０２ ２．２２	(Ｓ＝２) −３．６００ 最小	(Ｓ＝３) −０．７７１	——	——
２．０２ ２．２２ ３．０４	(Ｓ＝３) −２．２８４	——	——	

[0069] It is drawing 2 which made these the graph. The x axis of drawing 2 is the number s of the candidate of an outlier, and the y-axis is the value of Ut. When there were two or more values of Ut corresponding to s, the minimum thing in it was plotted. For example, although Ut was set to -0.318 and -0.514 in Table 3 in the case of the number s= 1 of the candidate of an outlier, it plotted using -0.514. The graph (1) shown with the alternate long and short dash line of drawing 2 shows the case where it is referred to as Ut=logsigma. The graph (2) shown by the dotted line shows the case where it is referred to as Ut=nlogsigma. The graph (3) shown as the solid line is the case where it is referred to as Ut=2s. The graph (4) expressed with the thick line is the value of nlogsigma and sigma+2s of Ut=nlog adding 2s.
[0070] A graph (2) shows that nlogsigma of the 1st term decreases as are mentioned above and the number s of the candidate of an outlier increases. And a graph (3) shows that 2s of the 2nd term increases. The value of sigma+2s of nlog(s) shown in a graph (4) is increasing, after taking the minimum value by s= 2. In sigma+2s of Ut=nlog, the minimum value is one piece. As shown in Table 5, it is the number s= 2 of the candidate of an outlier that a graph (4)

serves as the minimum, and it is the case where 2.02 and 2.22 are assumed as the outlier. This has obtained the same result also by the method of Kitagawa (GenshiroKitagawa(1979): "On the Use of Aic for the Detection of Outliers", Technometrics, Vol.21, No.2).

[0071] Next, example with one [another] more describes the relation between sigma+2s of Ut=nlog, and s. as data -1.40, -0.44, -0.30, -0.24, -0.22, -0.15, -0.13, 0.06, 0.10, 0.18, 0.20, 0.48, and 0. -- it is referred to as 63 and 1.01 The value of sigma+2s of Ut=nlog obtained by this is shown in Table 6.

[0072]

[Table 6]

外れ値の候補 大きい / 小さい	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	−9.42	−8.32	−6.24	−3.89
−1.40	−11.15	−11.12	−7.65	——
−1.40 −0.44	−8.82	−8.81	——	——
−1.40 −0.44 −0.30	−6.14	——	——	——

[0073] Based on Table 6, if a graph is written, it will become like drawing 3 . The x axis of drawing 3 is the number s of the candidate of an outlier, and the y-axis is the value of sigma+2s of Ut=nlog. The value of Ut is also large as s becomes large after the value of Ut becomes the minimum also in this case so that drawing 3 may show. Moreover, the number of outliers is considered to be a small number from the character of an outlier as compared with the total number of data. Therefore, when the calculation result of a detection statistic is henceforth shown in a table, only the data before and behind the minimum will be shown. It is also possible to detect an outlier, without specifying the number of the candidates of an outlier beforehand by using the character of this outlier. The value of a detection statistic also becomes large as are mentioned above and the number of outliers becomes large. Therefore, the calculation is terminated when the number of the candidates of an outlier computes a detection statistic in few order, it computes a detection statistic by increasing the number of the candidates of an outlier in order in not specifying the number of the candidates of an outlier, and the calculated detection statistic becomes large gradually. Thus, even if it is the case where the number of the candidates of an outlier is not specified beforehand, it becomes possible to detect an outlier. Therefore, when the number of the candidates of an outlier was beforehand specified that it mentioned above, the calculation result is large gradually and the bird clapper except became clear by comparing the calculation result of a detection statistic, without specifying the number of the candidates of an outlier by the system or the program, it becomes possible by stopping calculation of a detection statistic to detect an outlier. Next, in order to verify whether the detection statistic Ut is effective, an example 6 describes the thing in comparison with the result depended on the conventional calculation method from an example 2.

[0074] example 2. -- this example describes the outlier at the time of using Ut as a formula of a detection statistic using the data 2 of Grubbs All the data of Grubbs are quoted from the following reference.

"Procedures for detedting outying Observations in samples" Technometrics The data 2 of Vol.11 and 1-21Grubbs are the following value (the number of data is 12).

In 0.745, 1.832, 1.856, 1.884, 1.914, 1.916, 1.947, 1.949, 2.013, 2.023, 2.045, and the 2.327 original text, although 3 times of observed value and the averages have appeared, only the average is put on ascending order here. The calculation result of a detection statistic is shown in Table 7.

[0075]

[Table 7]

Ｇｒｕｂｂｓのデータ２と検出統計量

検定統計量 / 外れ値として除いたサンプル	$Ut=n\log\sigma+2s$
なし	-12.214
2.327	-9.578
2.045 2.327	-6.244
0.745	-20.497(2)
0.745 1.832	-16.610
0.745 2.327	-22.861(1)
0.745 1.832 2.327	-19.107
0.745 2.045 2.327	-19.145

[0076] That Ut takes the minimum value from Table 7 is the case where 0.745 and 2.327 are made into an outlier. The method of Kitagawa mentioned above has also brought the same result.
[0077] example 3. -- this example describes the case where the data 3 of Grubbs are used The data 3 of Grubbs are the following value (the number of data is 10).
The calculation result of 568, 570, 570, 570, 572, 572, 572, 578, 584, and a 596 detection statistic is shown in Table 8.
[0078]
[Table 8]

Ｇｒｕｂｂｓのデータ３（総データ数１０）と検出統計量

検定統計量 / 外れ値として除いたサンプル	$n\log\sigma+2s$
なし	21.109
596	15.975(2)
584 596	12.191(1)
568	21.075
568 570	21.158
568 596	16.319

[0079] That Ut takes the minimum value from Table 8 is the case where 584 and 596 are made into an outlier. According to Grubbs, 596 is made into the outlier. Dallas E. Johnson etc. made 584 and 596 the outlier. This is the same as the case where it asks by several 2. In addition, it is Dallas henceforth. When calling it E.Johnson etc., it shall be based on the following data.
Dallas E. Johnson Stephen A. McGuire and Geroge A. Milliken(1978): "Estimating sigma2 in the Presence of Outoliers" Technometrics Vol.20, No.4[0080] The case where what the same data were overlapped and made size double precision by example 4. and also the data of an example 3 is used is shown below. Data 4 are the following value (the number of data is 20).
568, 568, 570, 570, 570, 570, 570, 570, 572, 572, 572, 572, 572, 572, 578, 578, 584, 584, 596, 596 [0081] In this example, since size of data was made into double precision, the case where the number of the candidate of an outlier is set to 4 (s= 4) is explained. When the number of the candidate of an outlier is 4, a detection statistic will be computed to the following combination. That is, it is necessary to calculate the amount of statistics detection in case the number of the candidate of an outlier is 1 (s= 1), a detection statistic in case the number of the candidate of an outlier is 2 (s= 2), the detection statistic in the case of being the number 3 (s= 3) of the candidate of an outlier, and a detection statistic in case the number of the candidate of an outlier is 4 (s= 4). The detection statistic corresponding to the number s of the candidate of an outlier is as being shown below.
s= 1 Ut(-1) Ut(-20) s=2 Ut Ut (-1, -2) Ut(-19, -20) s=3 (-1, -20) Ut Ut (-1, -2, -3) Ut (-1, -2, -20) Ut (-1, -19, -20) (-18, -19, -20) s= 4 Ut(-1, -2, -3, -4) Ut(-1, -2, -3, -20) Ut(-1, -2, -19, -20) Ut(-1, -18, -19, -20) Ut (-17, -18, -19, -20) [0082] When the number of the candidate of an outlier is 4, it is possible to detect an outlier in the same order as a flow chart shown in drawing 1 . A different point is a point which computes each detection statistic from s= 1 to s= 4 which was mentioned above in Step 22 in drawing 1 . Thus, the calculation result of the computed detection statistic Ut is shown in Table 9.
[0083]

[Table 9]
Ｇｒｕｂｂｓのデータ４（総データ数２０）と検出統計量

検定統計量 / 外れ値として除いたサンプル	nlogσ+2s
なし	42.218
596	38.732
596 596	31.950
584 596 596	29.538(2)
584 584 596 596	24.382(1)
568	42.207
568 568	42.150
568 568 570	42.253
568 568 570 570	42.317
568 596	38.887
568 596 596	32.330
568 568 596 596	30.055
568 568 596	38.993
568 568 570 596	39.294

[0084] Outliers are 584, 584, 596, and 596 from Table 9. DallasE. Johnson etc. serves as same conclusion.
[0085] In addition, the number of the candidate of an outlier can be arbitrarily set up in the commonsense range based on the number of the inputted data. For example, when the number of the inputted data is 5 (N= 5), the number of the candidate of an outlier is a range with commonsense being referred to as 1 or 2 (s= 1 or 2). Moreover, if the number of the inputted data increases, it will not interfere with the part which also makes [many] the number of the candidate of an outlier. Thus, the number of the candidate of an outlier should be judged in the number of the inputted data, or its system according to the demand of the system of which grade to demand precision. In the example mentioned above or the example mentioned later, judgment whether to presume the number of outliers to be how many pieces shall be beforehand defined by the system, or shall be arbitrarily specified by the operator or the program.
[0086] The example using example 5., next the data of Rosner is described. This example is a case considered that size is comparatively as large as 54 and an outlier also exists mostly. Next, data are shown.
- 0.25, 0.68, 0.94, 1.15, 1.20, 1.26, 1.26, 1.34, 1.38, 1.43, 1.49, 1.49, 1.55, 1.56, 1.58, 1.65, 1.69, 1.70, 1.76, 1.77, 1.81, 1.91, 1.94, 1.96, 1.99, 2.06, 2.09, 2.10, 2.14, 2.15, 2.23, 2.24, 2.26, 2.35, 2.37, 2.40, 2.47, 2.54, 2.62, 2.64, 2.90, 2.92, 2.92, 2.93, 3.21, 3.26, 3.30, 3.59, 3.68, 4.30, 4.64, 5.34, 5.42, and 6.01 -- the data of this Rosner were taken from the following reference
Bernard Rosner(1977): "Percentage Point for a Generalized ESD Many-Outlier Procedure" Technometrics Vol.25, No.2, next the calculation result of Ut are shown in Table 10.
[0087]
[Table 10]

Ｒｏｓｎｅｒのデータと検出統計量

検出統計量 / 外れ値として除いたサンプル	nlogσ+2s
なし	8.564
-0.25	8.379
-0.25 0.68	9.607
6.01	5.415
5.42 6.01	3.008
5.34 5.42 6.01	-0.234
-0.25 6.01	4.939
-0.25 5.42 6.01	2.225
-0.25 0.68 5.42 6.01	3.346
4.64 5.34 5.42 6.01	-2.009
4.30 4.64 5.34 5.42 6.01	-3.305(1)
3.68 4.30 4.64 5.34 5.42 6.01	-2.995(2)
3.59 3.68 4.30 4.64 5.34 5.42 6.01	-2.681
3.30 3.59 3.68 4.30 4.64 5.34 5.42 6.01	-1.720
-0.25 5.34 5.42 6.01	-1.518

[0088] Rosner -- an outlier -- a maximum of -- it assumed that there was ten and authorized 5.34, 5.42, and 6.01 were made into the outlier at 5% of level of significance. Table 10 is a table showing the calculation result of the detection statistic by this example. When the calculation result of a detection statistic is shown in a table as having mentioned above, only the data before and behind the minimum value are shown. As shown in this table 10, when Formula Ut is used, outliers are 4.30, 4.65, 5.34, 5.42, and 6.01. In this case, although five outliers are detected as an outlier As Rosner assumed, even if it is the case where it is assumed that there are a maximum of ten outliers Or in the case of which [at the time of detecting an outlier automatically], a result detects these five outliers by measuring the detection statistic calculated whenever it increased the number of the candidates of an outlier, without specifying the number of the candidates of an outlier. In this example, although five outliers were detected, if an outlier carries out to to three pieces by this method, it is in agreement with Rosner.

[0089] The case where the example 6. normal random number, an index random number, and a uniform random number are made into sample data is described. Although, as for the normal random number and an index random number, an outlier may appear, I do not want an outlier to come from a uniform random number. Normal random number data are -2.666, -1.272, -0.042, and 0.140, 0.273, 0.415, 0.467, 1.160, 1.672 and 1.673 from (n= 10, -N (0 1)). The calculation result of Ut is shown in Table 11.

[0090]

[Table 11]

正規乱数と検出統計量データ(n=10, ～N(0,1)))

検出統計量 外れ値として除いたサンプル	nlogσ+2s
なし	2.294
1.167	3.475
1.672 1.673	4.998
-2.666	0.744(2)
-2.666 -1.272	0.436(1)
-2.666 1.673	2.320
1.160 1.672 1.673	6.546
-2.666 1.673	3.329
-2.666 -1.272 1.673	2.017
-2.666 -1.272 -0.042	2.559

[0091] Outliers are -2.666 and -1.272 from Table 11.
[0092] Next, the case where an index random number is used is shown. Data were taken from statistics mathematical" in the Takeuchi "phenomenon and action (**** Co.).
0.003, 0.021, 0.161, 0.178, 0.180, 0.210, 0.249, 0.413, 0.494, 0.562, 0.613, 0.879, 0.981, 1.059, 1.131, 1.264, 2.367, 3.669, 3.826, and the 4.193 total number of data are 20. The calculation result of Ut is shown in Table 12.
[0093]
[Table 12]

指数乱数と検出統計量

検出統計量 外れ値として除いたサンプル	nlogσ+2s
なし	5.041
4.193	3.894
4.193 3.826	2.074
4.193 3.826 3.669	-3.091(2)
4.193 3.826 3.669 2.367	-6.402(1)
0.003	6.889
0.003 0.021	8.706
0.003 0.021 0.161	10.573
0.003 0.021 0.161 0.178	12.405
4.193 0.003	5.896
0.003 4.193 3.826	4.261
0.003 4.193 3.826 3.669	-0.674
0.003 0.021 4.193	7.866
0.003 0.021 4.193 3.826	6.415

[0094] Outliers are 4.193, 3.826, 3.669, and 2.367 from Table 12.
[0095] Next, the case where the uniform random number of [0, 1] is used is described. The data are as follows by ten data.
0.283, 0.470, 0.643, 0.688, 0.916, 0.930, 0.945, 0.953, 0.973, and the calculation result of 0.995Ut are shown in Table 13.
[0096]
[Table 13]

[０，１] の一様乱数と検出統計量

外れ値として除いたサンプル ＼ 検出統計量	nlogσ+2s
なし	-14.484(1)
0.995	-11.003
0.995 0.973	-7.536
0.283	-13.649(2)
0.283 0.470	-12.533
0.283 0.995	-9.878
0.283 0.470	-4.109
0.283 0.995 0.973	-6.103
0.283 0.470 0.995	-8.435
0.283 0.470 0.643	-10.439

[0097] Since the case where there is no candidate of an outlier from Table 13 serves as the minimum, in the case of a uniform random number, there is no candidate of an outlier. Similarly about a uniform random number the candidate of an outlier did not have 1 case ******. It is a desirable result on the character of a uniform random number to have no candidate of an outlier.
[0098] About the data which have appeared in the example 7. Prior art, a detection statistic is calculated using Ut. Data are 5.71, 6.57, 7.29, 8.06, 10.00, and 15.00. A result becomes as it is shown in Table 14.
[0099]
[Table 14]

サンプルの組み合せと検出統計量の値

ケース	n	s	検出統計量
除かないとき	6	0	６.７７ $Ut_{(0)}$
５.７１を除いたとき	5	1	７.５５ $Ut_{(-1)}$
５.７１と１５.００ともに除いたとき	4	2	４.９９ $Ut_{(-1,-6)}$
１５.００を除いたとき	5	1	<u>３.９０</u> $Ut_{(-6)}$(最小値)

[0100] From Table 14, when 15.00 is removed, the minimum value and a bird clapper understand Ut and it makes 15.00 an outlier.
[0101] About the data of the masking effect stated with the example 8. conventional technology, a detection statistic is calculated using the formula shown in several 2. Data are 5.71, 6.57, 7.29, 8.06, 14.80, and 15.00. A result becomes as it is shown in Table 15.
[0102]
[Table 15]

サンプルの組み合せと検出統計量の値

ケース	n	s	検出統計量
除かないとき	6	0	8.06 $Ut_{(0)}$
5.71を除いたとき	5	1	8.61 $Ut_{(-1)}$
5.71, 15.00 を除いたとき	4	2	8.76 $Ut_{(-1,-6)}$
15.00を除いたとき	5	1	7.90 $Ut_{(-6)}$
15.00, 14.80 を除いたとき	4	2	3.44 $Ut_{(-6,-5)}$
5.71, 6.57 を除いたとき	4	2	9.15 $Ut_{(-1,-2)}$
5.71, 15.00, 14.80を除いたとき	3	3	4.51 $Ut_{(-1,-6,-5)}$
5.71, 6.57; 15.00 を除いたとき	3	3	9.65 $Ut_{(-1,-2,-6)}$

[0103] Table 15 -- Ut (-6, -5) the time -- the minimum value and a bird clapper -- understanding -- 15.00 and 14.80 -- an outlier -- carrying out . Thus, the masking effect is avoidable if the formula shown in several 2 is used.
[0104] Example 9. drawing 4 is drawing for explaining this example. The data processor of this example has measurement meanses, such as a sensor, and has an outlier detection means to detect an outlier, from the measured value obtained from this. Next, when there is an outlier, it has a data-processing means to calculate the average by the data except this. Next, the actual example of application is described. Five measured value obtained from a sensor is measured for every fixed time interval, an outlier is detected, and it considers making the average except this into measured value. x1 x2, x3, and x4 and x5 It is obtained as measurement data. The detection statistic Ut is calculated and an outlier is calculated. If there is an outlier, the average which does not have a bias except for an outlier can be calculated. Data are shown in Table 16 about time t1, t2, t3, and t4.
[0105]
[Table 16]

時刻	データ	平均値
t1	3.23, 1.24, 2.03, 2.86, 1.02	2.08
t2	2.06, 0.74, 2.19, 2.39, 2.73	2.01
t3	2.90, 0.08, 2.55, 2.23, 2.53	2.06
t4	1.83, 2.36, 3.36, 2.44, 1.95	2.39

[0106] The detection statistic Ut is calculated about the data of time t1. It is Ut (0) when not removing. = it is set to -0.711. the time of removing 1.02 -- Ut(-1) = -- it is set to Ut(-5) =0.709 when 0.952 and 3.23 are removed It is set to Ut (-1, -5) =2.760 when 1.02 and 3.23 are removed. It is as follows when this is summarized in Table 17. Ut (-1, -2) and Ut (-4, -5) Although it thinks, since it is uninfluential, in this example, it omits being shown in a table. That which is uninfluential similarly in the following tables will not display.
[0107]
[Table 17]

外れ値の候補 大きい方 小さい方	なし	3.23
なし	−0.711	0.709
1.02	0.952	2.760

[0108] Ut(s) in case there is no outlier are -0.711 and the minimum. Therefore, there is no outlier. Ut is calculated about the data of time t2, and it is shown in Table 18.
[0109]
[Table 18]

外れ値の候補 大きい方 小さい方	なし	2.73
なし	−1.969	0.205
0.74	−3.524	−2.504

[0110] Ut(s) when removing 0.74 are -3.524 and the minimum. Therefore, let 0.74 be an outlier. Ut is calculated about the data of time t3, and it is shown in Table 19.
[0111]
[Table 19]

外れ値の候補 大きい方 小さい方	なし	2.90
なし	0.057	2.112
0. 08	−3.753	−1.765

[0112] Therefore, let 0.08 be an outlier. Ut is calculated about the data of the following term t4, and it is shown in Table 20.
[0113]
[Table 20]

外れ値の候補 大きい方 小さい方	なし	3. 36
なし	−3.092	−3.388
1.36	−0.652	−0.616

[0114] Therefore, let 3.36 be an outlier. The data and the average which were updated become as it is shown in Table 21.
[0115]

[Table 21]

時刻	更新されたデータ					更新された平均値
t1	3.23,	1.24,	2.03,	2.86,	1.02	2.08
t2	2.06,	2.19,	2.39,	2.73		2.33
t3	2.90,	2.55,	2.23,	2.53		2.55
t4	1.83,	2.36,	2.44,	1.95		2.15

[0116] example 10. -- this example describes the data processor which has an outlier detection means and has an output means to tell an outlier Data as shown in Table 22 are obtained from an input process, and an outlier is detected by calculating Ut from this data. If the data of Table 22 are illustrated, it will become like drawing 5 .
[0117]
[Table 22]

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14
y	1	1.5	−1	−2	−1.5	0	−3	0	1.5	1	4	0	−1	0

[0118] If the detection statistic Ut is calculated about this data, it will become as it is shown in Table 23.
[0119]
[Table 23]

外れ値の候補 大きい方＼小さい方	なし	4	4, 1.5
なし	7.30	5.65	6.89
−3	7.39	<u>5.23</u>	6.54
−3, −2	8.22	5.77	6.54

[0120] Let 4 and -3 be outliers from Table 23. Conventionally, the graph as shows the data obtained from the input process to a display etc. at drawing 5 was displayed, after carrying out picking up of the value with which people are considered to be an outlier by viewing, it calculated, and it confirmed. The equipment by this invention can detect an outlier automatically, and it can check whether abnormalities have been in process environment and conditions.
[0121] Example 11. drawing 6 is drawing for explaining this example. The processing process which processes the data which inputted the outlier method of detection shown by drawing 6 between the input process of drawing 1 and the size judging process is added. In this example, data are received from the equipment which outputs weighted solidity with the inclination which increases with time, and the equipment which detects an outlier is stated from the data except the inclination which increases with time according to a processing process. Data are shown in Table 24. If this is illustrated, it will become like drawing 7 , and it turns out that data tend to increase with time.
[0122]
[Table 24]

t	1	2	3	4	5	6	7	8	9	10
y	2	4	3	6	6	5	6	11	6	9

[0123] Then, it is as follows when it asks for an inclination straight line by the least-squares method.
By lengthening the value of an inclination straight line from y=1.93+0.70t data, it becomes the following data.
- 0.63, 0.67, and - 1.03, 1.27, 0.57, -1.13, and -0. -- 83, 3.47, and - 2.23 and 0.07 -- if this is illustrated, it will become like drawing 8 If Ut is calculated from this amended data, it will become as it is shown in Table 25.
[0124]

[Table 25]

外れ値の候補 大きい方 / 小さい方	なし	3.47
なし	4.13	2.29
−2.23	4.92	2.63

[0125] Therefore, let y= 11 of t= 8 corresponding to the amended data 3.47 be an outlier. thus, the inclination which increases from data with the inclination which increases with time -- an amendment -- the outlier method of detection of an example 1 is applicable with things

[0126] example 12. -- when two or more weighted solidity is in one sample, this example calculates the diagonal element of a leverage (X(XT X)-1XT) in a processing process, and explains the equipment which detects an outlier based on this Data and the calculated leverage are shown in Table 26.

[0127]

[Table 26]

no.	x1	x2	x3	x4	テコ比
1	7	26	6	60	0.48
2	1	29	15	52	0.28
3	11	56	8	20	0.13
4	11	31	8	47	0.24
5	7	52	6	33	0.34
6	11	55	9	22	0.12
7	3	71	17	6	0.36
8	1	31	22	44	0.38
9	2	54	18	22	0.19
10	21	47	4	26	0.67
11	1	40	23	34	0.37
12	11	66	9	12	0.19
13	10	68	8	12	0.24
14	30	85	30	79	0.99

[0128] If the detection statistic Ut is calculated using a leverage, it will become as it is shown in Table 27.

[0129]

[Table 27]

外れ値の候補 大きい方 / 小さい方	なし	0.99	0.99 0.67
なし	−20.75	−22.99	−22.85
0.12	−17.37	−19.50	−19.39
0.12, 0.13	−14.02	−16.08	−16.12

[0130] - 22.99 is the minimum value. Therefore, let 0.99, sample no.14 [i.e.,], be an outlier. Since the influence which it has on the whole is large, the data with a large leverage of there being outlier detection equipment which can judge easily whether it is an outlier are effective.

[0131] example 13. -- when the inputted data are the model of regression analysis, this example searches for the remainder of regression analysis in a processing process, and explains the equipment which detects an outlier based on this The remainder searched for by data and regression becomes as it is shown in Table 28.

[0132]

[Table 28]

no.	x1	x2	x3	y	残差
1	7	26	60	78.5	0.346
2	1	29	52	74.3	1.545
3	11	56	20	104.3	-1.874
4	11	31	47	87.6	-1.783
5	7	52	33	95.9	-0.322
6	11	55	22	109.2	3.948
7	3	71	6	102.7	-1.339
8	1	31	44	72.5	-3.186
9	2	54	22	93.1	1.288
10	21	47	26	115.9	0.246
11	1	40	34	83.8	1.993
12	11	66	12	113.3	1.171
13	10	68	12	109.4	-2.033

[0133] If the detection statistic Ut is calculated using the remainder, it will become as it is shown in Table 29.
[0134]
[Table 29]

外れ値の候補 大きい方 / 小さい方	なし	3.948	3.948 1.993
なし	8.459	7.723	8.573
-3.186	8.762	7.873	8.816
-3.186, -2.033	9.746	8.939	10.030

[0135] Ut7.723 is the minimum. Therefore, the remainder 3.948, sample no.6 [i.e.,], serves as an outlier.
[0136] example 14. -- when the weighted solidity of the inputted data can be adapted in a canonical-correlation-analysis model with two or more, this example processes data, as a processing process is shown below, and explains the equipment which detects an outlier using this Data as shown in Table 30 are considered.
[0137]
[Table 30]

	特性値		要因			
no.	y1	y2	x1	x2	x3	x4
1	78.5	21.0	7	26	6	60
2	74.3	22.0	1	29	15	52
3	104.3	23.0	11	56	8	20
4	87.6	24.0	11	31	8	47
5	95.9	23.1	7	52	6	33
6	109.2	19.6	11	55	9	22
7	102.7	18.7	3	71	17	6
8	72.5	23.3	1	31	22	44
9	93.1	23.8	2	54	18	22
10	115.9	19.4	21	47	4	26
11	83.8	28.8	1	40	23	34
12	113.3	19.8	11	66	9	12
13	109.4	19.1	10	68	8	12

[0138] Canonical correlation analysis is performed to this data, and two synthetic variate functions of y1 and y2 can be found. A synthetic variate function value can be found using this synthetic variate function. A leverage is calculated based on a synthetic variate function value. A leverage is as follows.
0.27, 0.31, 0.18, 0.12, 0.10, 0.20, 0.19, 0.30, 0.12, 0.30, 0.63, and 0. -- 13 and 0.16 -- if the detection statistic Ut is calculated about this leverage, it will become as it is shown in Table 31
[0139]
[Table 31]

外れ値 大きい方 / 小さい方	なし	0.63
なし	−26.00	−29.03
0.10	−22.03	−24.96

[0140] Therefore, let a leverage 0.63 be an outlier. This is the sample of no.11.
[0141] example 15. -- weighted solidity is classified into two groups, the inputted data perform discriminant analysis according to two or more factors in a processing process, and this example is described about the equipment which detects the outlier in each group based on this processed data Data are shown in Table 32.
[0142]
[Table 32]

データ

	no.	x1	x2	判別関数値
グループ1	1	6	0	−3.80
	2	0	2	−0.56
	3	0	3	−0.84
	4	1	2	−1.20
	5	1	3	−1.48
	6	1	4	−1.76
グループ2	7	4	0	−2.53
	8	4	1	−2.82
	9	5	0	−3.17
	10	5	1	−3.45
	11	5	2	−3.73
	12	0	4	−1.12

[0143] If data are plotted, it will become like drawing 9 . Into a group 1, sample no.1 seems to be [sample no.12] an outlier from drawing at a group 2. If discriminant analysis is carried out about data and it asks for a discriminant function, it will become y=-0.634*x1-0.281*x2. In the case of no.1, it will be set to discriminant-function value =-0.634*6-0.281*0=-3.80 if a discriminant-function value is calculated using this discriminant function. The discriminant-function value was put on the right column of Table 32. If Ut is calculated about the discriminant-function value of a group 1, it will become as it is shown in Table 33.
[0144]
[Table 33]

外れ値 大きい方 / 小さい方	なし	−0.56
なし	0.33	2.19
−3.80	−2.24	−0.33

[0145] Therefore, let the discriminant-function value -3.80 and sample no.1 be outliers. If Ut is calculated about the discriminant-function value of a group 2, it will become as it is shown in Table 34.
[0146]
[Table 34]

外れ値 大きい方 / 小さい方	なし	−1.124
なし	−1.00	−2.24
−3.73	0.94	−0.24

[0147] Therefore, let the discriminant-function value -1.12 and sample no.12 be outliers.
[0148] example 16. -- weighted solidity is classified into three groups, the inputted data perform discriminant analysis

according to two or more factors in a processing process, and this example is described about the equipment which detects the outlier in each group based on this data Data are shown in Table 35.

[0149]

[Table 35]

データ	no.	x1	x2	判別関数値 f_1	f_2	ユークリッド距離 d
グループ1	1	0	3	3.45	-1.21	3.65
	2	0	4	4.60	-1.62	4.88
	3	1	2	1.46	-1.33	1.97
	4	1	3	2.61	-1.74	3.14
	5	1	4	3.76	-2.14	4.33
グループ2	6	4	0	-3.39	-2.11	3.99
	7	5	0	-4.23	-2.63	4.98
	8	4	1	-2.24	-2.51	3.36
	9	5	1	-3.08	-3.04	4.33
	10	4	2	-1.08	-2.91	3.10
	11	5	2	-1.93	-3.44	3.94
グループ3	12	4	5	2.37	-4.13	4.76
	13	4	6	3.52	-4.53	5.74
	14	5	5	1.52	-4.65	4.89
	15	5	6	2.67	-5.06	5.72
	16	6	5	0.67	-5.18	5.22
	17	6	6	1.83	-5.58	5.87
	18	0	2	2.30	-0.81	2.44

[0150] If discriminant analysis is carried out about data and a discriminant-function value is calculated, it will become like the right-hand side of Table 35, and the column of a discriminant-function value. 2 sets of discriminant-function values are acquired by 3 group ****'s. Generally a discriminant-function value is ********** (group number -1). Based on the following formula, Euclidean distance is found from 2 sets of these discriminant-function values, and this is shown in the column of the Euclidean distance of Table 35.

Euclidean distance [of a discriminant-function value] d (=(f12+f22) 1/2)

Ut is as follows when Ut is calculated for every group about the Euclidean distance of a discriminant-function value. Ut of a group 1 is shown in Table 36.

[0151]

[Table 36]

グループ1

外れ値 大きい方 / 小さい方	なし	4.88
なし	0.02	1.41
1.97	0.34	1.84

[0152] The case where Ut does not have the candidate of an outlier is the minimum. Therefore, there is no outlier. Ut of a group 2 is shown in Table 37.

[0153]

[Table 37]

グループ2

外れ値 大きい方 / 小さい方	なし	4.98
なし	-2.92	-2.01
3.10	-1.16	-0.22

[0154] The case where Ut does not have the candidate of an outlier is the minimum. Therefore, there is no outlier. Ut of

a group 3 is shown in Table 38.
[0155]
[Table 38]
グループ3

外れ値 大きい方 / 小さい方	なし	5.87	5.87 5.74
なし	0.67	2.66	4.62
2.44	-3.10	-0.49	2.03
2.44 , 4.76	-0.94	1.88	4.77

[0156] Therefore, Euclidean distance makes 2.44 and sample no.18 an outlier. In addition, although this example described the case where weighted solidity was three groups, it can carry out similarly about three or more groups.
[0157] example 17. -- this example describes the case where a detection statistic is calculated in a formula which is different in several 2 shown in the above-mentioned example The formula of the detection statistic Uta uses several 3.
[0158]
[Equation 3]

$$Uta = n \log \alpha + \frac{b_2}{2} + 2S$$

ただし n は サンプル数

S は 外れ値の候補の個数

α は $x_1, \ldots, x_n$ をサンプルデータとすると

$$\alpha^2 = \frac{\Sigma(x_i - \bar{x})^2}{n}$$

$$b_2 = \frac{\Sigma(x_i - \bar{x})^4}{\alpha^2}$$

[0159] as data -- (the total 15 data), -1.40, -0.44, -0.30, and -0. -- 24, -0.22, -0.13, -0.15, 0.06, and 0. -- 10, 0.18, 0.20, 0.39, 0.48, and 0. -- 63 and 1.01 are used A calculation result becomes as it is shown in Table 39.
[0160]
[Table 39]

除いたサンプル	$n \log \sigma + b_2/2 + 2s$	参考 $n \log \sigma + 2s$
なし	-7.267	-9.419
1.01	-5.860	-8.323
-1.40	-9.770	-11.153 (1)
1.01, -1.40	-10.101 (1)	-11.117
1.01, 0.63	-3.577	-6.237
-1.40, -0.44	-7.421	-8.815
-1.40,-0.44,1.01	-7.825	-8.806
-1.40,1.01,0.63	-8.575	-9.561

[0161] Therefore, let 1.01 and -1.40 be outliers from a table. Since -1.40 is made into the outlier to use sigma+2s of Ut=nlog so that it may turn out that the column of reference of Table 39 is seen, many one outlier is specified in Uta. However, by many other data, the same result as the case where several 2 formula is used has been obtained. Several 3 formula has the feature of detecting many one outlier, depending on [this] the case. In addition, b2 / 2 of the 2nd term of several 3 considered the index of amendment of merit that the normal distribution of "Takeuchi" is applied as reference.
Statistic Ts which expresses the appropriateness of a normal distribution model according to Takeuchi (Takeuchi ** (1976): "criteria of distribution [of an information statistic], and appropriateness of model", and "mathematical-

science" NO.153 Saiensu-Sha Co., Ltd., 12-18) (following Takeuchi's statistic) It is zi about the data of measurement size n, and z in zi (i = 1, ..., n). It is as follows as an average.
the Ts =-log sigma-b 2 / 2n -- here -- sigma2 ={sigma (zi-z) 2} / nb2 = {sigma (zi-z) 4} / nsigma4 -- it is close to such a suitable normal distribution model that the value of this Takeuchi's statistic is large
[0162] When there are two, the normal distribution model a and the regular part model b, and distribution of the normal distribution model a shows a value smaller than distribution of the normal distribution model b, the direction of the normal distribution model a is Data zi from the normal distribution model b. Average = many values near 0 are shown. It is called the correction term and has b2 in the 2nd term of the formula which calculates above-mentioned Takeuchi's statistic / 2n of amendment meanings for the value of logsigma in the 1st term. Therefore, the value of logsigma which Takeuchi's statistic has in the 1st term influences greatly. Therefore, characterization of this Takeuchi's statistic is made with the value of Distribution sigma. Therefore, the value of Takeuchi's statistic becomes large, so that distribution is small, and the pattern (namely, small pattern of distribution) near the normal distribution model a will be shown from the normal distribution model b, so that the value of this Takeuchi's statistic is large.
[0163] The example 46 is creating the formula for calculating a detection statistic from this example 18 below example 18. on the basis of the regression-analysis explanation variable selection criterion shown in drawing 10 . Several 2 mentioned above and several 3 are considered on the basis of AIC. AIC is an example of a regression-analysis explanation variable selection criterion. therefore, the point which can do so the same effect as the example which mentioned it above when an example 46 calculated a detection statistic from the following examples 18 on the basis of regression-analysis explanation variable selection criterions other than a AIC **** regression-analysis explanation variable selection criterion is boiled, attached and explained It thinks on the basis of the regression-analysis explanation variable selection criterion which shows a-32 number to drawing 10 from four formulas of a detection statistic shown by the example 46 from an example 18, and these formulas are roughly divided and can be classified into two types. The 1st group is the same form even as the example mentioned above, and is the nlogsigma+ 2nd term (the + 3rd term).
It comes out. The 2nd group is a multiplication type and is adjustment factor xsigma. moreover -- as data -1.40, -0.44, -0.30, -0.24, -0.22, -0.15, -0.13, 0.06, 0.10, 0.18, 0.20, 0.39, 0.48, and 0. -- 63 and 1.01 are used in all the examples after this example If this is illustrated, it will become like drawing 11 . The number of the outliers which can be found by the difference in a formula differs to the same data. Therefore, it uses according to the application which does not need to take out many outliers, and business to take out many outliers, being able to choose a formula.
[0164] The formula of the detection statistic St is shown in several 4.
[0165]
[Equation 4]
$$St = \frac{(n-2)(n-1)}{(n-S-2)(n-S-1)} \sigma$$

[0166] If St is calculated, it will become as it is shown in Table 40.
[0167]
[Table 40]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.534	0.565	0.667
−1.40	0.462	0.458	0.537
−1.40 −0.44	0.547	0.572	0.752

[0168] Therefore, let 1.01 be an outlier to -1.40.
[0169] The formula of the example 19. detection statistic Ft is shown in several 5.
[0170]
[Equation 5]
$$Ft = \frac{n+S}{n-S} \sigma$$

[0171] If Ft is calculated, it will become as it is shown in Table 41.

[0172]
[Table 41]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.53	0.55	0.62
−1.40	0.45	0.43	0.46
−1.40 −0.44	0.51	0.49	0.54

[0173] Therefore, let 1.01 be an outlier to -1.40.
[0174] The formula of the example 20. detection statistic Tt is shown in several 6.
[0175]
[Equation 6]

$$Tt = \frac{n^2 - n - S - 2}{n(n-S-2)(n-S-1)}\alpha$$

[0176] If Tt is calculated, it will become as it is shown in Table 42.
[0177]
[Table 42]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.041	0.046	0.059
−1.40	0.038	0.041	0.052
−1.40 −0.44	0.048	0.055	0.079

[0178] Therefore, let -1.40 be an outlier.
[0179] The formula of the example 21. detection statistic TIt is shown in several 7.
[0180]
[Equation 7]

$$TIt = n \log \alpha + 2S + \frac{S(S-2)}{n}$$

[0181] If TIt is calculated, it will become as it is shown in Table 43.
[0182]
[Table 43]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	−9.42	−8.39	−6.24
−1.40	−11.22	−11.12	−9.31
−1.40 −0.44	−8.82	−8.56	−6.49

[0183] Therefore, let -1.40 be an outlier.
[0184] The formula of the example 22. detection statistic Wt is shown number 8.
[0185]
[Equation 8]

$$Wt = S \log n + n \log \alpha$$

[0186] If Wt is calculated, it will become as it is shown in Table 44.
[0187]
[Table 44]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	-9.42	-7.78	-5.11
-1.40	-10.51	-9.99	-8.11
-1.40 -0.44	-7.69	-7.35	-5.63

[0188] Therefore, let -1.40 be an outlier.
[0189] The formula of the example 23. detection statistic Pt is shown number 9.
[0190]
[Equation 9]

$$Pt = \frac{1}{(n - S - 1)^2} \alpha$$

[0191] If Pt is calculated, it will become as it is shown in Table 45.
[0192]
[Table 45]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.002722	0.0033	0.0045
-1.40	0.002714	0.0031	0.0043
-1.40 -0.44	0.00373	0.0045	0.0069

[0193] Therefore, let -1.40 be an outlier.
[0194] An example 24. detection statistic hematocrit formula is shown number 10.
[0195]
[Equation 10]

$$Ht = \frac{(n-1)(n+S+1)}{(n+1)(n-S-2)} \alpha$$

[0196] If hematocrit is calculated, it will become as it is shown in Table 46.
[0197]
[Table 46]

外れ値の 大きい方 候補 小さい方	なし	1. 01	1. 01 0. 63
なし	147.14	135.69	135.89
－1. 40	110.85	93.36	89.37
－1. 40 －0. 44	111.45	95.18	96.28

[0198] Therefore, let -1.40, and 1.01 and 0.63 be outliers.
[0199] The formula of a U.1 t example 25. detection statistic is shown number 11.
[0200]
[Equation 11]

$$U.1t = n \log \alpha + 2 * S * 1.1$$

[0201] U. If 1 t is calculated, it will become as it is shown in Table 47.
[0202]
[Table 47]

外れ値の 大きい方 候補 小さい方	なし	1. 01	1. 01 0. 63
なし	-9.42	-8.32	-6.24
－1. 40	-11.15	-11.12	-9.56
－1. 40 －0. 44	-8.82	-8.81	-7.22

[0203] Therefore, let -1.40 be an outlier.
[0204] The formula of a U.9 t example 26. detection statistic is shown number 12.
[0205]
[Equation 12]

$$U.9t = n \log \alpha + 2 * S * 0.9$$

[0206] U. If 9 t is calculated, it will become as it is shown in Table 48.
[0207]
[Table 48]

外れ値の 大きい方 候補 小さい方	なし	1.01	1.01 0.63
なし	-9.42	-8.52	-6.64
－1.40	-11.35	-11.52	-10.16
－1.40 －0.44	-9.22	-9.41	-8.02

[0208] Therefore, let 1.01 be an outlier to -1.40.
[0209] The formula of the example 27. detection statistic Uut is shown number 13.
[0210]
[Equation 13]

$$U_{ut} = \frac{n-1}{n-S} \alpha$$

[0211] If Uut is calculated, it will become as it is shown in Table 49.
[0212]
[Table 49]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.498	0.478	0.496
−1.40	0.391	0.341	0.334
−1.40 −0.44	0.407	0.356	0.358

[0213] Therefore, let -1.40, and 1.01 and 0.63 be outliers.
[0214] The formula of the example 28. detection statistic Bt is shown number 14.
[0215]
[Equation 14]

$$Bt = \frac{n+1}{n-s-1}\alpha$$

[0216] If Bt is calculated, it will become as it is shown in Table 50.
[0217]
[Table 50]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.610	0.598	0.637
−1.40	0.489	0.438	0.444
−1.40,-0.44	0.522	0.473	0.501

[0218] Therefore, let 1.01 be an outlier to -1.40.
[0219] The formula of the example 29. detection statistic Dt is shown number 15.
[0220]
[Equation 15]

$$Dt = \frac{(n+1)(n+s+1)}{n-s}\alpha$$

[0221] If Dt is calculated, it will become as it is shown in Table 51.
[0222]
[Table 51]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	8.539	8.243	8.604	9.360
−1.40	6.734	5.911	5.833	6.063
−1.40,-0.44	7.056	6.212	6.304	6.954

[0223] Therefore, let -1.40, and 1.01 and 0.63 be outliers.
[0224] The formula of the example 30. detection statistic Gt is shown number 16.
[0225]
[Equation 16]

$$Gt = \frac{n+S+1}{n-S+1}\alpha$$

[0226] If Gt is calculated, it will become as it is shown in Table 52.
[0227]
[Table 52]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.534	0.547	0.607	0.702
−1.40	0.447	0.417	0.438	0.482
−1.40,-0.44	0.498	0.466	0.501	0.580

[0228] Therefore, let 1.01 be an outlier to -1.40.
[0229] The formula of the example 31. detection statistic Qt is shown number 17.
[0230]
[Equation 17]

$$Qt = \frac{1}{(n-S)^2}\alpha$$

[0231] If Qt is calculated, it will become as it is shown in Table 53.
[0232]
[Table 53]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.00237	0.0028	0.0038	0.0054
−1.40	0.00231	0.00258	0.0034	0.0049
−1.40,-0.44	0.0031	0.00359	0.0051	0.0087

[0233] Therefore, let -1.40 be an outlier.
[0234] The formula of the example 32. detection statistic It is shown number 18.
[0235]
[Equation 18]

$$It = \frac{n+S}{(n-S)^2}\alpha$$

[0236] If It is calculated, it will become as it is shown in Table 54.
[0237]
[Table 54]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.036	0.043	0.056	0.081
−1.40	0.035	0.039	0.051	0.074
−1.40,-0.44	0.046	0.054	0.077	0.130

[0238] Therefore, let -1.40 be an outlier.
[0239] The formula of the example 33. detection statistic Vt is shown number 19.

[0240]
[Equation 19]

$$Vt = \frac{n-1}{(n-S)^2}\sigma$$

[0241] If Vt is calculated, it will become as it is shown in Table 55.
[0242]
[Table 55]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.033	0.037	0.045	0.060
−1.40	0.030	0.031	0.037	0.049
−1.40,-0.44	0.037	0.040	0.051	0.078

[0243] Therefore, let -1.40 be an outlier.
[0244] The formula of the example 34. detection statistic Et is shown number 20.
[0245]
[Equation 20]

$$Et = \frac{n+1}{(n-S-1)^2}\sigma$$

[0246] If Et is calculated, it will become as it is shown in Table 56.
[0247]
[Table 56]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.038	0.043	0.055	0.075
−1.40	0.035	0.038	0.047	0.067
−1.40,-0.44	0.045	0.050	0.070	0.122

[0248] Therefore, let -1.40 be an outlier.
[0249] The formula of the example 35. detection statistic Jt is shown number 21.
[0250]
[Equation 21]

$$Jt = \frac{n(n+S+1)}{(n-S)^2}\sigma$$

[0251] If Jt is calculated, it will become as it is shown in Table 57.
[0252]
[Table 57]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.569	0.634	0.782	1.040
−1.40	0.518	0.537	0.648	0.866
−1.40,-0.44	0.641	0.690	0.901	1.391

[0253] Therefore, let -1.40 be an outlier.
[0254] The formula of the example 36. detection statistic Vtd is shown number 22.
[0255]
[Equation 22]

$$Vtd = n \log \sigma - \frac{b_2}{2} + 2 * S$$

[0256] If Vtd is calculated, it will become as it is shown in Table 58.
[0257]
[Table 58]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	-11.57	-10.79	-8.90
−1.40	-12.54	-12.13	-10.55
−1.40,-0.44	-10.21	-9.79	-8.15

[0258] Therefore, let -1.40 be an outlier.
[0259] The formula of the example 37. detection statistic BBt is shown number 23.
[0260]
[Equation 23]

$$BBt = \frac{n+1}{(n-S-1)^2} \sigma$$

[0261] If BBt is calculated, it will become as it is shown in Table 59.
[0262]
[Table 59]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.044	0.050	0.064
−1.40	0.041	0.044	0.056
−1.40,-0.44	0.052	0.059	0.084

[0263] Therefore, let -1.40 be an outlier.
[0264] The formula of the example 38. detection statistic CCt is shown number 24.
[0265]
[Equation 24]

$$CCt = \frac{n(n-S+1)}{(n-S)^2} \sigma$$

[0266] If CCt is calculated, it will become as it is shown in Table 60.
[0267]
[Table 60]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.569	0.559	0.587
−1.40	0.453	0.403	0.405
−1.40,-0.44	0.481	0.431	0.450

[0268] Therefore, let 1.01 be an outlier to -1.40.
[0269] The formula of the example 39. detection statistic DDt is shown number 25.
[0270]
[Equation 25]

$$DDt = \frac{n(n+S+1)}{(n-S)^2}\sigma$$

[0271] If DDt is calculated, it will become as it is shown in Table 61.
[0272]
[Table 61]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.589	0.634	0.782
−1.40	0.518	0.537	0.648
−1.40,-0.44	0.641	0.690	0.901

[0273] Therefore, let -1.40 be an outlier.
[0274] The formula of the example 40. detection statistic GGt is shown number 26.
[0275]
[Equation 26]

$$GGt = \frac{n+S+1}{(n-S+1)^2}\sigma$$

[0276] If GGt is calculated, it will become as it is shown in Table 62.
[0277]
[Table 62]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.033	0.039	0.051
−1.40	0.032	0.035	0.044
−1.40,-0.44	0.042	0.047	0.063

[0278] Therefore, let -1.40 be an outlier.
[0279] The formula of the example 41. detection statistic Upt is shown number 27.
[0280]
[Equation 27]

$$Upt = \frac{n+S-1}{(n-S+1)^2}\sigma$$

[0281] If Upt is calculated, it will become as it is shown in Table 63.
[0282]
[Table 63]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.467	0.478	0.531
−1.40	0.391	0.365	0.383
−1.40,-0.44	0.435	0.408	0.439

[0283] Therefore, let 1.01 be an outlier to -1.40.
[0284] The formula of the example 42. detection statistic Zt is shown number 28.
[0285]
[Equation 28]

$$Zt = \frac{n}{n-s}\sigma$$

[0286] If Zt is calculated, it will become as it is shown in Table 64.
[0287]
[Table 64]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.534	0.515	0.538	0.585
−1.40	0.421	0.369	0.365	0.379
−1.40,-0.44	0.441	0.388	0.394	0.435

[0288] Therefore, let -1.40, and 1.01 and 0.63 be outliers. Zt detects more outliers.
[0289] The formula of the example 43. detection statistic Kt is shown number 29.
[0290]
[Equation 29]

$$Kt = \frac{1}{n-s}\sigma$$

[0291] If Kt is calculated, it will become as it is shown in Table 65.
[0292]
[Table 65]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63	1.01 0.63 0.48
なし	0.036	0.037	0.041	0.049
−1.40	0.030	0.028	0.030	0.034
−1.40,-0.44	0.034	0.032	0.036	0.043

[0293] Therefore, let 1.01 be an outlier to -1.40.
[0294] The formula of the example 44. detection statistic Xt is shown number 30.
[0295]
[Equation 30]

$$Xt = \frac{1}{(n-S)(n+S)} \alpha$$

[0296] If Xt is calculated, it will become as it is shown in Table 66.
[0297]
[Table 66]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	0.0024	0.0025	0.0028
−1.40	0.0020	0.0019	0.0020
−1.40,-0.44	0.0023	0.0022	0.0024

[0298] Therefore, let 1.01 be an outlier to -1.40.
[0299] The formula of the example 45. detection statistic HQt is shown number 31.
[0300]
[Equation 31]

$$HQt = n \log \alpha + C * S * \log (\log(n))$$
$$C > 2$$

[0301] If HQt is calculated, it will become as it is shown in Table 67.
[0302]
[Table 67]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	-9.42	-7.41	-4.59
−1.40	-10.24	-9.47	-7.40
−1.40,-0.44	-7.16	-6.61	-4.72

[0303] Therefore, let -1.40 be an outlier.
[0304] The formula of the example 46. detection statistic AICt is shown number 32.
[0305]
[Equation 32]

$$AICt = n \log \alpha^2 + 2 * S$$

[0306] If AICt is calculated, it will become as it is shown in Table 68.
[0307]
[Table 68]

外れ値の候補 大きい方 / 小さい方	なし	1.01	1.01 0.63
なし	-17.80	-17.61	-15.43
−1.40	-23.27	-25.19	-24.08
−1.40,-0.44	-20.59	-22.57	-21.39

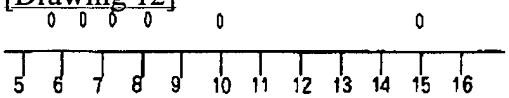
[0308] Therefore, let 1.01 be an outlier to -1.40.
[0309]
[Effect of the Invention] Since an outlier is detected based on the detection statistic computed when inputting the value according to the 1st invention, it is not necessary to consider size comparison of a table as calculated value like before. Moreover, it is not necessary to set up the number of an outlier beforehand. Or what is necessary is just to specify a number to detect as an outlier of maximums. moreover, the number of an outlier -- or since it is not necessary to change a calculation method by the outlier of the larger one, and the outlier of the smaller one, computation is easy, and there is also little computational complexity and it ends Moreover, by the conventional method, although the outlier might be undetectable with the masking effect, since this is avoidable, a more exact result is obtained. Moreover, when an outlier does not exist, it judges with not existing.
[0310] According to the 2nd invention, the value for computing an outlier can be chosen easily.
[0311] According to the 3rd invention, since an outlier is calculated only by simple comparison of a detection statistic, processing becomes easy.
[0312] Since according to the formula in the 4th invention a detection statistic will serve as the minimum if the value from which it separated most is removed when there is an outlier, an outlier can be calculated using this.
[0313] According to the formula in the 5th invention, since there is a correction term, an outlier can be calculated more exactly.
[0314] Since distribution is included in the 1st item, if the value from which it has separated most is removed according to the formula in the 6th invention, the property in which distribution becomes small can be used.
[0315] According to the formula in the 7th invention, since the number of the candidate of an outlier is included in the 2nd item, the downward tendency of the 1st item by having increased the number of an outlier can be offset.
[0316] According to the formula in invention of the octavus, the augend of the 2nd item can be adjusted by carrying out the multiplication of the coefficient to the 2nd item.
[0317] According to the formula in the 9th invention, the downward tendency of distribution can be adjusted with a coefficient.
[0318] According to the formula in the 10th invention, the formula of regression analysis can be applied and an outlier can be detected.
[0319] According to the 11th invention, when there is a processing process, the outlier of data various type is detectable.

[0320] According to the 12th invention, an outlier is detectable by processing the value for which it does not depend on time from the value depending on time, either.
[0321] According to the 13th invention, even if it is the case where two or more weighted solidity exists in one sample, an outlier can be calculated by calculating a leverage.
[0322] If it is the value which can apply the technique of regression analysis according to the 14th invention, the outlier of this remainder can be calculated by searching for the remainder of regression analysis.
[0323] According to the 15th invention, an outlier can be calculated even when a canonical-correlation-analysis model can be applied.
[0324] An outlier can be calculated, when according to the 16th invention weighted solidity is classified into two or more groups and discriminant analysis can be performed.
[0325] According to the 17th invention, the data processor which can detect an outlier easily can be obtained by using the above outlier methods of detection. By examining an environmental condition when this outlier is obtained, new knowledge and information also become acquired.
[0326] According to the 18th invention, by using the above outlier methods of detection, an outlier can be detected easily and the data processor which can be removed can be obtained. Since the outlier is removed, reliability of the result obtained by this data processor is improving.

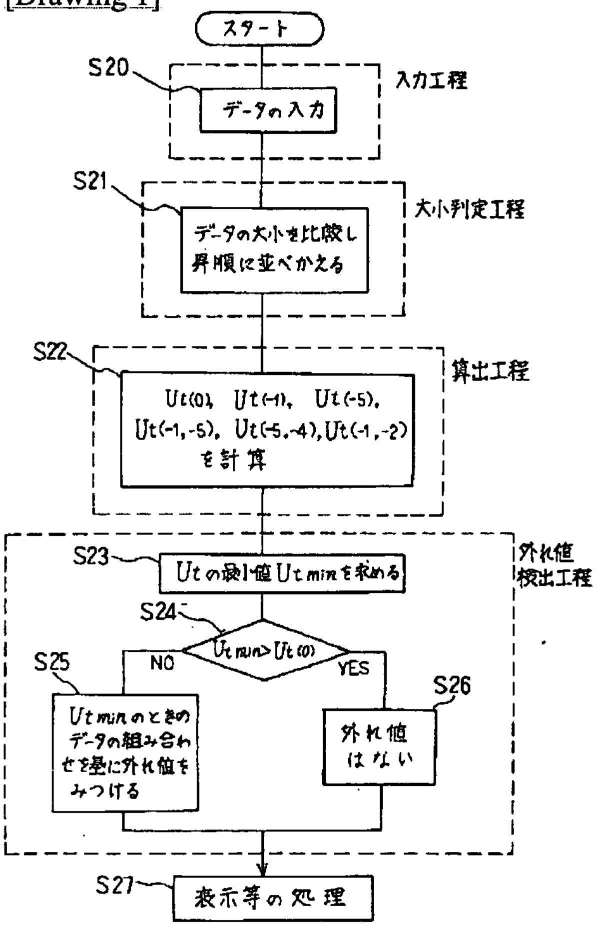
[Translation done.]

DRAWINGS

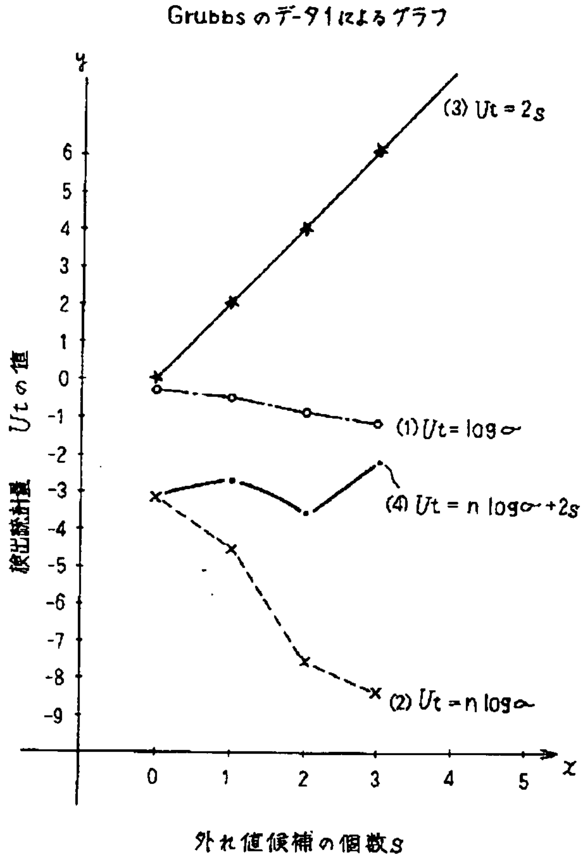
[Drawing 12]



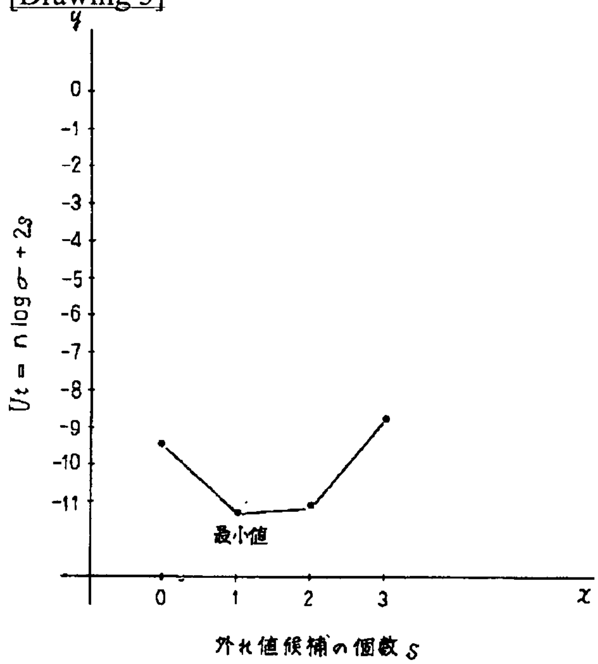
[Drawing 1]



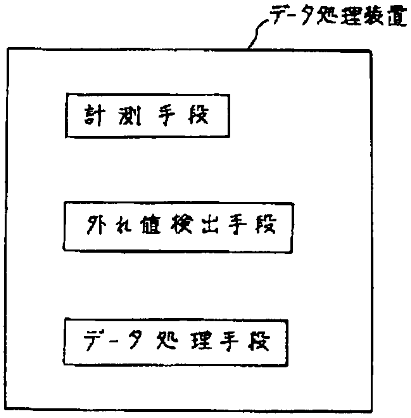
[Drawing 2]



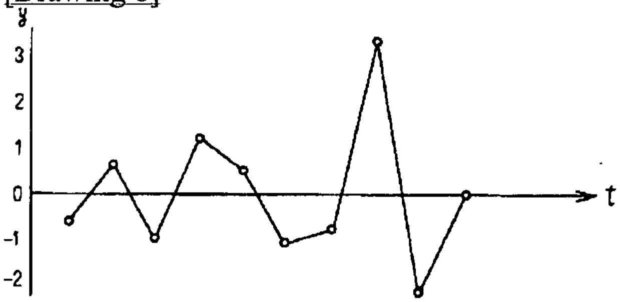
[Drawing 3]



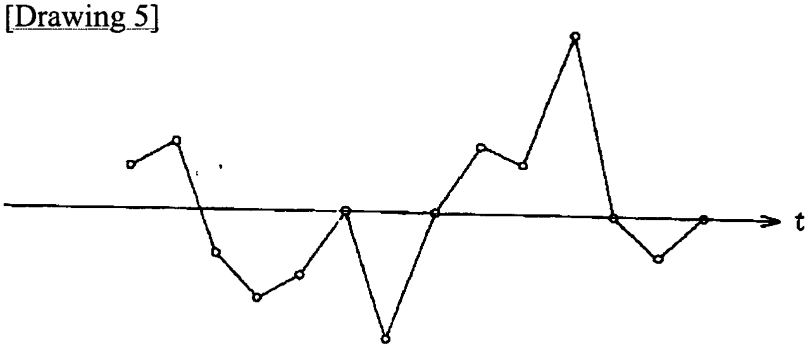
[Drawing 4]



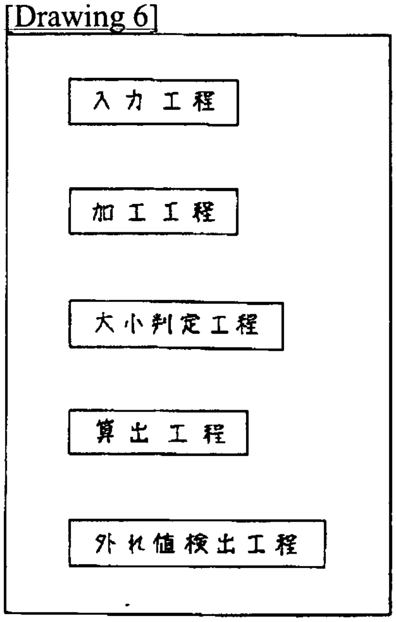
[Drawing 8]



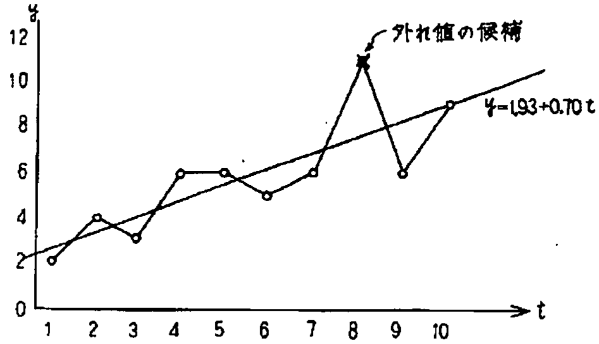
[Drawing 5]



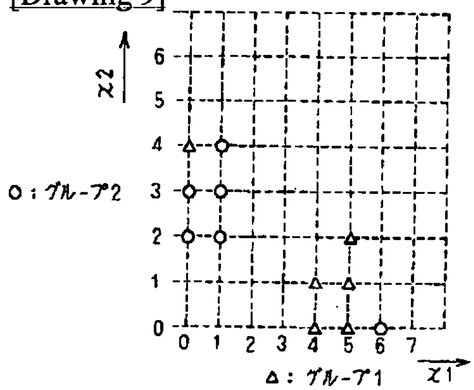
[Drawing 6]



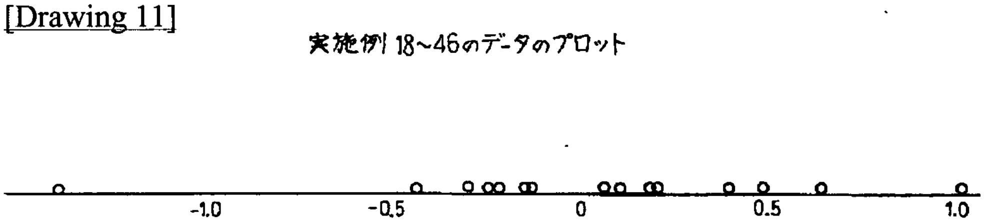
[Drawing 7]



[Drawing 9]



[Drawing 11]



[Drawing 10]

回帰分析説明変数選択規準

1 $\underline{aic} = n \log(\Sigma e_i^2) + 2P$

n：データの個数
P：説明変数の個数
R：重相関係数

2 $\underline{tic} = aic + P(P-2)/n$

3 $\overline{\text{佐和}} = 1 - \frac{(n-2)(n-1)}{(n-P-2)(n-P-1)}(1-R^2)$

4 $\overline{\text{芳賀他}} = 1 - \frac{(n-1)(n+P+1)}{(n+1)(n-P-1)}(1-R^2)$

5 $\underline{\text{竹内}} = \frac{n^2-n-P-2}{n(n-P-2)(n-P-1)}\sigma^2$　　σ^2：残差の分散

6 $\underline{\text{予測}} = \frac{(n+1)(n-2)}{n(n-P-2)}\sigma^2$

7 $\underline{Sp} = \frac{\Sigma e_i^2}{(n-P-1)(n-P-2)}$　　Σe^2：残差の2乗和

8 $\overline{\text{上田}} = 1 - \frac{(n+P+1)}{(n-P-1)}(1-R^2)$

9 $\underline{aicC} = n \log(\Sigma e_i^2) + C \cdot P$　　$C > 0$

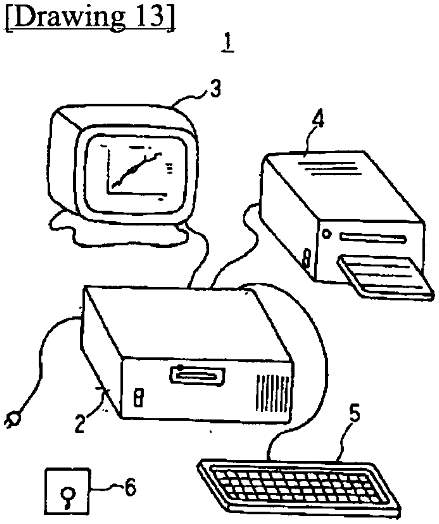
10 $\underline{HQ} = n \log(\Sigma e_i^2) + C \cdot P \cdot \log(\log(n))$　$C > 2$

11 $\underline{fpe} = \frac{n+P}{n-P}\Sigma e_i^2$

12 $\underline{Helms} = \frac{P}{n(n-P)}\Sigma e_i^2$

13 $\underline{\text{シュワルツ}} = P \log n + n \log \sigma$

[Drawing 13]

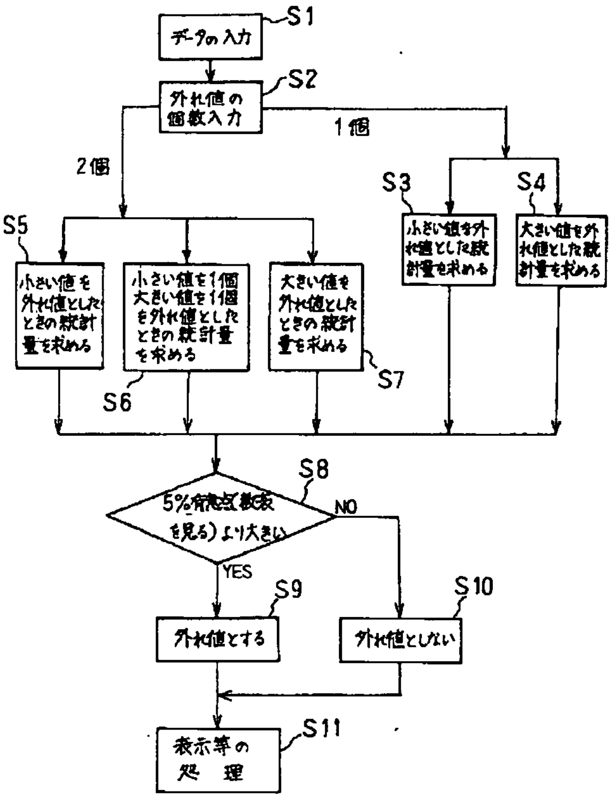


1 ： 情報処理装置
2 ： コンピュータ（FDD付き）
3 ： ディスプレイ
4 ： プリンタ
5 ： キーボード
6 ： フロッピーディスク

[Drawing 14]

従来方式

危険率5%(有意水準5%)の場合で説明する。外れ値の個数が1あるいは2個のときとする。



[Translation done.]